e420400



#2

OIPE

RAW SEQUENCE LISTING DATE: 01/18/2002 PATENT APPLICATION: US/10/027,806 TIME: 09:17:38

Input Set: N:\Crf3\RULE60\10027806.raw
Output Set: N:\CRF3\01182002\J027806.raw

```
1 <110> APPLICANT: Swanson, Ronald V.
              Feldman, Robert A.
              Schleper, Christa
      4 <120> TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
      5 <130> FILE REFERENCE: DCORP.002A
      6 <140> CURRENT APPLICATION NUMBER: US/10/027,806
      7 <141> CURRENT FILING DATE: 2001-12-21
      8 <150> PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
W--> 9 <151> PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
     10 <160> NUMBER OF SEO ID NOS: 123
     11 <170> SOFTWARE: FastSEO for Windows Version 3.0
     13 <210> SEO ID NO: 1
     14 <211> LENGTH: 32998
                                                          ENTERED
     15 <212> TYPE: DNA
     16 <213> ORGANISM: Cenarchaeum symbiosum
     17 <220> FEATURE:
     18 <221> NAME/KEY: CDS
     19 <222> LOCATION: (7604)...(8908)
     20 <221> NAME/KEY: CDS
     21 <222> LOCATION: (8961)...(9767)
     22 <221> NAME/KEY: CDS
     23 <222> LOCATION: (10545)...(10922)
     24 <221> NAME/KEY: CDS
     25 <222> LOCATION: (13944)...(14612)
     26 <221> NAME/KEY: CDS
     27 <222> LOCATION: (18638)...(20149)
     28 <221> NAME/KEY: CDS
     29 <222> LOCATION: (20554)...(20955)
     30 <221> NAME/KEY: CDS
     31 <222> LOCATION: (20956)...(21834)
     32 <221> NAME/KEY: CDS
     33 <222> LOCATION: (25151)...(26377)
     34 <221> NAME/KEY: CDS
     35 <222> LOCATION: (27535)...(28002)
     36 <221> NAME/KEY: CDS
     37 <222> LOCATION: (28065)...(29483)
     38 <400> SEQUENCE: 1
     39
              gatectique etetgegett attgeageea tggactquee ggeegtgegg ggetaaataa
                                                                                     60
     40
              agetgaggeg eegeetgeag getetgetea geegttgatt atacagtaet egeactegea
                                                                                    120
     41
              ggtcttgctt tcgtcatctt ttgccctgca cttggcgtgc tggcccgact ggcactggac
                                                                                    180
     42
              gcagatctcg tctatgttgt ctatcgaggt gtactttttg gactggtcaa agccgaagcc
                                                                                    240
```

gccgtccttt ggtccaacca tgaagcggtc agggccgtta tgccttaatt atctttgccg

teggggeggg geogettetg eeggegggag eeceeettga ggeegeteee eggetegttg

43

44

300

360

RAW SEQUENCE LISTING DATE: 01/18/2002 PATENT APPLICATION: US/10/027,806 TIME: 09:17:38

45			cctgatcata				420
46	cccactatca	gtataaacag	cctagccacg	cccccattc	tgcccatgcg	tgtaatatgc	480
47	tgcacctgta	aacaaacatt	gcgcggggca	tgggccgtcc	ggacagacag	aactgcccat	540
48			ggtaagctac				600
49			taatcatctg				660
50			ttgccagaaa				720
51	caactgcccg	atcgtttttc	aggtacctac	tgggcttttt	ggagttcgtt	gatcaaaagc	780
52			gccagtcttg				840
53			tgcttttcac				900
54			tgtagttgag				960
55			tccttctctt				1020
56			gtttgattat				1080
57			acggtatttg				1140
58			taccggggga				1200
59	tgcatccacc	gtggacccat	gagacaacga	gccggcagcg	gtgcacagca	ccgagtacaa	1260
60			acgacctgcc				1320
61			tgagaattct				1380
62			tgattcatga				1440
63			actgacagca				1500
64	gatcagttac	ttttaagcca	atccttttac	attcttctct	cgatatgcgc	cttccatgac	1560
65			gacattattt				1620
66	cagactcgtt	ctcaaacatg	tatctagcca	gccatttttg	tacaagttcc	acacttagct	1680
67	tctggctgct	aatgcatttt	tgaattagcc	cgggaggata	ttgccctaac	agtggaagcc	1740
68	atgcgccgag	cctccccggg	tgtttttctg	acaccgactg	tgcttcttga	aactcgctaa	1800
69			atgtgcatgc				1860
70	cggtgggacc	tatcgatgag	tgtttaccca	ttaccaggca	attcgatgag	catgcaagca	1920
71	tcgcagctgc	cgacatcgcg	gcatatggga	taatgatccg	gacattttta	aattttgcac	1980
72			tcggtggact				2040
73	aatctaattt	cttagtcttt	aaatcacgca	tcatcctcat	aaatccatac	aggtcaccat	2100
74	ttgttatgag	agcttcatta	gacgtatgcg	cttcgtccgt	tatccagttg	gtcgcataca	2160
75			tactgttgca				2220
76	caggtgcctg	ctcaccgact	gactccatta	gtcgaagtac	gtcgtttgct	atttctgcgt	2280
77	aactaggcat	ttatgtagtt	gaaatgacta	cccgcgggaa	tcataccata	gtctgtgtcg	2340
78	tatgacttgc	cttttttctt	catcaatttc	tcatattcct	catgcagttc	gagcaggatg	2400
79	gccatgtcac	tacgtttgtt	cgtctgtttt	gttgtctcgg	gcttttggtc	catattatta	2460
80	tccatgctag	taaaggacta	tgttccttta	aaaaggttcg	tgattttaat	ttccaagtgt	2520
81	ttgcctcgca	atttcctcca	aggcacatga	aaaacgggcc	acaggcagag	cacagcatcc	2580
82	gctggggacc	catgaaataa	gcccccggcg	gtgcacagca	tccgctgggg	gctcaataaa	2640
83			gtctctatgt				2700
84	gtaaaggctg	gctcaccact	cgccgaagct	tgtgggatac	accaccttcc	tatcaacgca	2760
85	gtcttcttct	gcgaaccttc	atccgaagaa	ggaatatctt	gtctcgggat	aggattcgtg	2820
86			agcctagatg				2880
87			gcctctctgt				2940
88			gcagaggccg				3000
89			agcagcctca				3060
90	_		ccaaaccgcg				3120
91			ttttctgtca				3180
92			gtctatgaat				3240
93			agcggatttc				3300
	,,,,,,			-	-		

RAW SEQUENCE LISTING DATE: 01/18/2002 PATENT APPLICATION: US/10/027,806 TIME: 09:17:38

							2260
94			tgccgcccca ·				3360
95			ctgcaggaaa				3420
96			tgactcccag				3480
97			ccacggggtc				3540
98			caccgggttg				3600
99			tacccgacaa				3660
100			ccttagctcg				3720
101			catacccttt				3780
102						cagctgcagg	3840
103			caggactaat				3900
104						gaacttgcag	3960
105			tcatggtctc				4020
106			gttctcgtca				4080
107						aaacgctccc	4140
108			actggtttcc				4200
109	aggatcgact	aactccaggc	tgacgacgca	ttgcctggaa	acccttgcgc	ttacggtggt	4260
110	gcggattctc	accgcactat	gctgttactg	ccaccaggat	ctgcaataga	aatcggtcca	4320
111			tcgtcccaat				4380
112	cacggtgcac	gtctggagta	tcggtactct	gctttagccc	cgtccgtttt	tgtggcgccc	4440
113	tcgctcggca	ggtaagttgt	tacacacttt	ttgaaggata	gctacttctg	agcttacctc	4500
114	cctgctgtct	tggcgatgac	acgcactttg	gcttgacact	tagcagaaat	ttggggacct	4560
115	taactccagt	ctgggttaaa	cccctctcgg	tcgtgaacct	tacgtcacac	gaacccgtgt	4620
116						atctcttccc	4680
117	cgcgccaccc	tatcagtgct	ctaccggaaa	caccatctcc	acatagcacg	ccctgcgaga	4740
118	cgcttcggtt	ggaactagca	agcgccagtc	tagattggtt	tttgacccct	attcccaagt	4800
119	cacacaaacg	agttgcacgt	cagaactgct	gcagacctcc	agtgggcttt	cgcccacctt	4860
120						taacgcactt	4920
121						cgcctttcta	4980
122						aacgcatgac	5040
123						ccgaaaaaat	5100
124			_			acccccttc	5160
125						agagatattt	5220
126		-		-		ctactcgggt	5280
127						cggaacattt	5340
128						catctcggcc	5400
129	-					acttgggaaa	5460
130						gttcgacctc	5520
131						ggtgcgtgca	5580
132						ctagcaatcc	5640
133	_	•	cggcatattc				5700
134			gggggccgct				5760
135						ggtgatccga	5820
136						cctcaagttc	5880
137			acctcactaa				5940
138			attcactgcg				6000
139			cageceteag				6060
140			cccattgtca				6120
141			gccgtggccc				6180
142	CCGCTAATTC	gooccacigo	cccygagagc	aatyytyyca	actayaggca	aggatctcgc	6240



RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/027,806

DATE: 01/18/2002
TIME: 09:17:38

							6200
143	tcgttacctg	acttaacagg	acatctcacg	gcacgagctg	gcgacggcca	tgcaccacct	6300
144	ctcagcttgt	ctggtagagt	cttcagcttg	accttcacac	tgctgtctct	ccgggtaaga	6360
145	tttctggcgt	tgactccaat	tgaaccgcag	gcttcacccc	ttgtggtgct	ccccgccaa	6420
146	ttcctttaag	tatcatactt	gcgtacgtac	ttcccaggcg	gcaaacttaa	cggcttccct	6480
147	gcggcactgc	actggctctt	acgccaatgc	atcactgagt	ttgcattgtt	tacagctggg	6540
148	actacccqqq	tatctaatcc	ggtttgctcc	cccagctttc	atccctcacc	gtcggacgtg	6600
149	ttctagtaga	ccgccttcgc	cacagggggt	catcgataga	tcagaggatt	ttaccccttc	6660
150	ctaccgagta	ccgtctacct	ctcccactcc	ctagccgtgc	agtatttccg	gcagcctatg	6720
151	cattgagcgc	atagatttaa	ccgaaaactt	acacggcagg	ctacggatgc	tttaggccca	6780
152	ataatcctcc	tgaccacttg	aggtgctggt	tttaccgcgg	cggctgacac	cagaacttgc	6840
153	ccacccctta	ttcqccqqtq	gttttaagac	cggtaaaaga	tttctttagc	agaaaacact	6900
154	cggattaacc	ttatcatact	ttcgcacatt	gcaaagtttt	ctcgcctgct	gcgccccata	6960
155	agacctagat	ccgtgtctca	gtacccatct	ccgggcctct	cctctcagag	cccgtatctg	7020
156	ttatagcctt	ggtgggccat	tacctcacca	acaagctgat	agaccgcagt	cccatcctac	7080
157	ggcgataaat	catttgggcc	acaaaccatt	ccaggcatag	tggcctatcg	gatattattc	7140
158	tcagtttccc	gaggttatcc	ccgtccatag	gttagattga	ctacgtgtta	ctgagccgtc	7200
159	tgccttgtat	tgctacaatg	actcgcatgg	cttagtatca	atccgatagc	agtcaggtcc	7260
160	ggcaggatca	accggattca	taattggatt	attttttt	tgttaagtac	gcttgtactt	7320
161	ttggaattga	acagaatgca	cataatcttc	acatctcaga	tatgaccctt	cgatcatacc	7380
162	ctcattctgt	gtgcgtaact	ggaggccagc	gaatcacaat	atggtacaat	accatgcatt	7440
163	catcgcaagc	gccgctcttg	cgtcacgtac	gatcggatcg	gcccgtccat	gggcatataa	7500
164	accategeeg	atttccqccc	ccggcagccc	cgatcagggg	ccggatctgc	ctgtatgatg	7560
165	gcgatccgcc	ctgattaaat	tatgggggga	gcggcctgct	gccgtggatc	tggaacgcga	7620
166	gtacagggca	aagaccggcg	actcaaccca	gatctttgcc	aggtcgaaaa	agtaccacgt	7680
167	caacaaaatc	agccacaaca	taaggttcta	cgagccgtat	ccgtttgtga	caaggtccgc	7740
168	gageggaag	cacctcgtcg	acgtggacgg	gaacaagtat	gtagactact	ggatggggca	7800
169	ctagageeta	atactggggc	acgcgccggc	gccagtcagg	toggcagtag	aggggcagct	7860
170	teaceacaae	tagatecaca	ggaccgtcaa	cgagcagacg	atgaatctct	cggagataat	7920
171	acacaacaca	gtaagcgtgg	cagaaaagac	aaggtacgtc	acgtcgggga	cggaggccgt	7980
172	catgtatgcg	gcaaggctgg	cacacacaca	tacqqqcaqa	aaaataatag	caaaggcgga	8040
173	caacaactaa	cacagataca	catcagaact	gctcaagtcg	gtcaactggc	cgtatgatgt	8100
174	acccasasac	agaaaactca	togacgaaga	gcactctata	tccattccgt	acaacgatct	8160
175	tgaaggttcc	ctagatatte	ttagacacac	aggcgacgac	ttggcatgcg	tgataatcga	8220
176	accactacta	aacaacaaca	getgeatace	ggcggatgag	gactatctgc	gcggcataca	8280
177	geogetgetg	cattcaagg	acacactact	tatectegae	gagatagtga	cagggttccg	8340
178	atttaaattt	gactacacat	atoctocaoc	agggctggac	cccgatatag	tggcgctcgg	8400
179	caagatagte	gggggggggat	tececatagg	ggtgatatgc	ggcaaggacg	aggtgatgga	8460
180	aatotooaac	actatatogo	atgcaaagtc	cgacagggcg	tacatcggcg	gcggcacatt	8520
181	ctctccaac	cccaccacaa	tgacagcggg	cacaacaaca	ctcggggagc	tcaaaaagag	8580
182	aaaggggaga	atatacccga	ggataaactc	catgggggac	gacgcaaggg	acaagctctc	8640
183	aaagatattt	gggaacaggg	tatecataac	cggaaggggc	tcqctqttca	tgactcactt	8700
184	tattaaaaat	adedecades	gggtctcaaa	tgctgcagat	gcggcagcct	gcgatgttga	8760
185	getgetgeac	aggtaccacc	tggacatgat	cacccgggac	ggcatattct	ttctgccggg	8820
186	getgetgeac	gocatatoro	caacacactc	aaaggccgac	ctcaagacca	tgtattccgc	8880
	caayetgggg	tttggagaag	goctatgagg	tataggooguo	gaggaaactt	tgattatacg	8940
187	accayaycyc	coorgagaay	atgatactet	teggeaagag	casccccacc	gagctggtgc	9000
188	ggegtgetge	cotootataa	accaacaacc	agttcaggg	ggcaataggc	ctgtacggga	9060
189	gccaygcgga	ggaggaggg	cadaacadd	acatectaes	caaaaaaaaa	ctggcccaga	9120
190	aadtcctcaa	. gyacyacccy	cayaacayyy	catactttas	ccaactacta	gagcttgaca	9180
191	acagggcaaa	aaaytactct	gatycyatca	. cycyclicga			3200

RAW SEQUENCE LISTING DATE: 01/18/2002 PATENT APPLICATION: US/10/027,806 TIME: 09:17:38

				+	~~~~~~	~~~~~~~~	9240
192	acaaggacgc	gcccgcgtac	aacaacaagg	ccatageeea	ggeegagete	ggagacacgg	9300
193	catccgcgct	ggaaaactac	ggcagggcca	tegaggeega	ceegeggtae	gegeeggege	9360
194	gcttcaacag	ggccgtgctg	ctcgacaggc	tgggcgagca	Lgaggaggeg	anggatag	9420
195	tcgacagggc	agccgagctg	gaccgacgca	agccgaaccc	gaggilelae	aayyyyacay	9480
196	tgctcggcaa	gatgggcagg	cacgaagagg	cgctggcctg	cttcaagggc	gracitaga	9540
197	ggcatcccgg	ccacgccgac	tcacagttcc	acgtggggat	agagettace	gagerrygea	9600
198	ggcacgccga	ggccctcggg	gagettgeat	cactgcccgc	ggagcaccgc	gagaacycca	9660
199	atgtattgta	tgccagggcg	cgcagcctct	cgggccttgg	cagggaggac	gaatecatag	9720
200	cgcacctgca	aaaggcggcc	aaaaaagatt	ccaagacgat	aaaaaagtgg	geeegegeag	9720 9780
201	aaaaggcctt	tgacggaata	cgggacgatc	ccggttcaaa	aagatageeg	getagaggat	9840
202	cttttttctt	gccgcgtcaa	tccgcatcat	gcggaccttt	tttttgggcc	ccacaagtcg	
203	cgattcatag	actggtacat	agaccacctc	caccgccttt	gcggcaaact	cctcccgcag	9900 9960
204	gtcgcgcatg	ccgtcaggcg	ggggcccgcg	cagcttctct	tttagttttg	agagegeete	
205	ttctgtctcc	acctcggggc	tccgcacatt	ctctgacgca	tegagtatee	teegegggta	10020
206	cggctccacc	gegeegggee	ccgtcttgta	gggaaagtcc	gtctcgccgc	cgtgccggtc	10080
207	aaggcacatc	atcccttctg	attccgcaaa	gacatgctct	tctagctcga	ggtcgaccct	10140
208	gttcttgccg	agccctgccg	agagcgtctt	gtgtatgcgc	gacttggacc	ttatgggaaa	10200
209	gacgccgtcg	cctagcacca	cctcgatcac	gttctggtcc	accttgatcg	ggtgaaccgc	10260
210	ctttctgaaa	aaatccgcag	agtacctggc	ggagacccgg	atcagcgcct	cgtggaccag	10320
211	ctttacagaa	tgcacatgga	cgtcttcttt	ccgcggggcc	ctcataaggg	ccctaaaggc	10380
212	acccgtcttc	tttgcctcta	tcatggcccg	agccgactcc	tcagtcatgg	cgttccgcag	10440
213	gaccgccgtc	ctggtctttc	cagtcatccc	ctgccgcacc	ccgcataagg	catactatac	10500
214	aacgcaaggc	aaggtaataa	tagcctgccg	tctgtaacgg	ccgtatgagg	tcggagggca	10560
215	ggcccggata	catcgaaaag	ttcctaaaga	gggcggacaa	ggcgatagac	aatgcagtcg	10620
216	agcagggcgt	caagagggca	gacgagatac	tagatgacgc	agtcgagctc	ggcaagatca	10680
217	ccatagacaa	qqcqcaaaaa	agaagcgatg	tgctgctcaa	gcaggccgag	cgggagagca	10740
218	agcggctcaa	gtcaaggggc	gccaaaaagc	tcgaaaaggg	cataggggcg	gcaaaaaaga	10800
219	tggcagccgg	caaqqqcqac	gcgctagaga	ccctggcaaa	gctcggcgag	ctgagaaagg	10860
220	cggggatcat	aacqqaqaaq	gagtttcgcg	ccaagaaaaa	gaagcttctc	gcggagatct	10920
221	gacttgaage	cactagacta	tacccgggac	ggctcgataa	aggaggtcac	aaagaggtgg	10980
222	tttataggca	caccatccct	tgtcgacctt	qcaggcgagc	tcggcatatc	tgagagcaac	11040
223	atattccacg	tgacatttcc	cgacggcgca	aagaccaccc	tgcatacgca	cgagggcggg	11100
224	cagetgetea	tagtgacctc	gggcaccggc	agcatgtcaa	tatttgaaaa	gaccggcgga	11160
225	aacasaacaa	aatttgcaat	aaaagagaca	gacaggatac	cgctaaagca	gggcagcatc	11220
226	cagtagatag	ctaccaacat	gctcacgtgc	acqqcqcaac	agacggcacc	accetgteec	11280
227	atatageggt	aaactacccg	tegecategg	gaaaggagcc	gtatacatta	tggtatgaat	11340
228	ccgactttgc	cagccgggtc	accggcgtgc	totaaattat	attatttgag	cctctccagt	11400
229	atcgacagge	ttacaaggtt	ggtcatcgtt	atccccttqc	ggatcacttc	ccttctagtc	11460
230	ttttcacagt	acttgttgac	gctctggtgg	ctcttttcgc	tggatacctc	gagcaccaca	11520
231	atattcccac	tatacaggaa	ggcaaacttt	ggcggaatcc	togagoatac	atccatgcgc	11580
232	cotattcccq	ccctacctat	cttctccttt	cttgccacta	ccgacgccac	ctcqcqtatq	11640
232	tectecteg	aatctccgta	ttccagatag	tacatggata	catageteat	cccgggggat	11700
234	tocctttcca	atateteete	gtccatgctg	aataaatagg	aagggccgcc	cccacaaccc	11760
234	tocacogoct	ttatcattt	ggggccgttt	ttgaacettg	ccagcaggta	ataactaacc	11820
235	tantantan	autacuuctt	gcttttggac	gagaaagtgg	cggtcacaaa	gtacateteg	11880
	agtagettee	togaccacca	ccattcctcc	ttgaggtgt	togaggeatg	atcatacata	11940
237	tatagagetag	aggostatos	ccccggggag	acctccatct	tagacataca	taggateege	12000
238	catygegige	ayycatatcc	tccatgccgc	cttagagaga	adadacccca	cctataacca	12060
239	ggaaccggtt	aalalclagt	ggat cast ac	cccggggggc	taaaactacc	ataccastcs	12120
240	geeeegggge	aggegtgegt	ggatccatgc	yatayıtatı	cadaaccagg	acyccyacoa	1212V



VERIFICATION SUMMARY

PATENT APPLICATION: US/10/027,806

DATE: 01/18/2002

TIME: 09:17:39

Input Set : N:\Crf3\RULE60\10027806.raw
Output Set: N:\CRF3\01182002\J027806.raw

L:9 M:256 W: Invalid Numeric Header Field, Wrong Prior FILING DATE:YYYY-MM-DD

L:7288 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:120